Claims

1. A zinc finger polypeptide library in which each polypeptide comprises more than one zinc finger and wherein each polypeptide has been at least partially randomised such that the randomisation extends to cover the overlap of a single pair of zinc fingers.

2. A library according to claim 1 wherein each polypeptide comprises between three and six zinc fingers.

3. A library according to claim 1 or claim 2, wherein one and a half zinc fingers are randomised in each polypeptide.

4. A set of zinc finger polypeptide libraries which encode overlapping zinc finger polypeptides, according to any one of claims 1 to 4, wherein the polypeptides may be assembled after selection to form a multifinger zinc finger polypeptide.

5. A set according to claim 4, comprising a pair of libraries encoding three-zinc finger polypeptides.

6. A library or set of libraries according to any preceding claim, wherein the randomised positions are selected from positions -1, 1, 2, 3, 5 and 6.

7. A library according to any preceding claim, wherein the randomisation of amino acid residues is restricted such that the following amino acids may appear at the given positions:

Position	Possible Amino Acids
-1	R, Q, H, N, D, A, T
1	S, R, K, N
2	D, A, R, Q, H, K, S, N
3	H, N, S, T, V, A, D
5	I, T, K
6	R, Q, V, A, E, K, N, T

R, Q, V, A, E, K, N, T

8. A set of two libraries according to claim 7 for selecting a three-finger zinc finger protein, wherein the following amino acids may appear at the given positions:

Library 1		Library 2	
F1:	**amino acid**	**F1:**	**amino acid**
-1	R, Q, H, N, D, A		
2	D, A, R, Q, H, K, S, N		
3	H, N, S, T, V, A, D		
5	I, T		
6	R, Q, V, A, E, K, N, T		
F2			
-1	R, Q, H, N, D, A, T		
1	S, R		
2	D, A, R, Q, H, K, S, N		
3	H, N, S, T, V, A, D	3	H, N, S, T, V, A, D
		6	R, Q, V, A, E, K, N, T
F3			
		-1	R, Q, H, N, D, A, T
		1	R, K, S, N
		2	D, A, R, Q, H, K, S, N
		3	H, N, S, T, V, A, D
		5	K, I, T
		6	R, Q, V, A, E, K, N, T

9. A library according to claim 1, wherein the amino acids at positions -1, 2, 3 and 6 are selected as follows:

a) if base 4 in the quadruplet is G, then position +6 in the α-helix is Arg or Lys;

b) if base 4 in the quadruplet is A, then position +6 in the α-helix is Glu, Asn or Val;

c) if base 4 in the quadruplet is T, then position +6 in the α-helix is Ser, Thr, Val or Lys;

d) if base 4 in the quadruplet is C, then position +6 in the α-helix is Ser, Thr, Val, Ala, Glu or Asn;

e) if base 3 in the quadruplet is G, then position +3 in the α-helix is His;

f) if base 3 in the quadruplet is A, then position +3 in the α-helix is Asn;

g) if base 3 in the quadruplet is T, then position +3 in the α-helix is Ala, Ser or Val; provided that if it is Ala, then one of the residues at -1 or +6 is a small residue;

h) if base 3 in the quadruplet is C, then position +3 in the α-helix is Ser, Asp, Glu, Leu, Thr or Val;

i) if base 2 in the quadruplet is G, then position -1 in the α-helix is Arg;

j) if base 2 in the quadruplet is A, then position -1 in the α-helix is Gln;

k) if base 2 in the quadruplet is T, then position -1 in the α-helix is His or Thr;

l) if base 2 in the quadruplet is C, then position -1 in the α-helix is Asp or His.

m) if base 1 in the quadruplet is G, then position +2 is Glu;

n) if base 1 in the quadruplet is A, then position +2 Arg or Gln;

o) if base 1 in the quadruplet is C, then position +2 is Asn, Gln, Arg, His or Lys;

if base 1 in the quadruplet is T, then position +2 is Ser or Thr.

10. A library according to any preceding claim, wherein each zinc finger has the general primary structure

(A) X^a C X_{2-4} C X_{2-3} F X^c X X X X L X X H X X^b H - linker

-1 1 2 3 4 5 6 7 8 9

wherein X (including X^a, X^b and X^c) is any amino acid.

11. A library according to claim 10 wherein X^a is $^F/_Y$-X or P-$^F/_Y$-X.

12. A library according to claim 10 or claim 11 wherein X_{2-4} is selected from any one of: S-X, E-X, K-X, T-X, P-X and R-X.

13. A library according to any one of claims 10 to 12 wherein X^b is T or I.

14. A library according to any one of claims 10 to 13 wherein X_{2-3} is G-K-A, G-K-C, G-K-S, G-K-G, M-R-N or M-R.

15. A library according to any one of claims 10 to 14 wherein the linker is T-G-E-K or T-G-E-K-P.

16. A library according to any one of claims 10 to 15 wherein position +9 is R or K.

17. A library according to any one of claims 10 to 16 wherein positions +1, +5 and +8 are not occupied by any one of the hydrophobic amino acids, F, W or Y.

18. A library according to claim 17 wherein positions +1, +5 and +8 are occupied by the residues K, T and Q respectively.

19. A method for preparing a library of nucleic acid binding proteins of the Cys2-His2 zinc finger class capable of binding to a target nucleic acid sequence, comprising the steps of:

a) selecting a model zinc finger polypeptide from the group consisting of naturally occurring zinc finger polypeptides and consensus zinc finger polypeptides; and

b) randomising more than one finger therein according to any one of claims 1 to 9.

20. A method according to claim 19, wherein the model zinc finger is a consensus zinc finger whose structure is selected from the group consisting of the consensus

structure P Y K C P E C G K S F S Q K S D L V K H Q R T H T G, and the consensus structure P Y K C S E C G K A F S Q K S N L T R H Q R I H T G E K P.

21. A method according to claim 19 wherein the model zinc finger is a naturally occurring zinc finger whose structure is selected from one finger of a protein selected from the group consisting of Zif 268 (Elrod-Erickson *et al.*, (1996) Structure 4:1171-1180), GLI (Pavletich and Pabo, (1993) Science 261:1701-1707), Tramtrack (Fairall *et al.*, (1993) Nature 366:483-487) and YY1 (Houbaviy *et al.*, (1996) PNAS (USA) 93:13577-13582).

22. A method according to claim 21 wherein the model zinc finger is finger 2 of Zif 268.

23. A method for determining the presence of a target nucleic acid molecule, comprising the steps of:

a) preparing a nucleic acid binding protein by the method of any preceding claim which is specific for the target nucleic acid molecule;

b) exposing a test system comprising the target nucleic acid molecule to the nucleic acid binding protein under conditions which promote binding, and removing any nucleic acid binding protein which remains unbound;

c) detecting the presence of the nucleic acid binding protein in the test system.

24. A method according to claim 23, wherein the presence of the nucleic acid binding protein in the test system is detected by means of an antibody.

25. A method according to claim 23 or claim 24 wherein the nucleic acid binding protein, in use, is displayed on the surface of a filamentous bacteriophage and the presence of the nucleic acid binding protein is detected by detecting the bacteriophage or a component thereof.